Measurement error is the error caused in observation and it makes the bias of coefficients. Many method are trying to overcome this, but these can't be applied if there are no sufficient data. For example, fixed effect or random effect is a hidden bias which we can't catch it. It come from the individual owns trait or circumstance which effect on data. We can reduce these bias by averaging and this is effective for fixed effect. Random effect can be erased very well if we use proper distribution for it, which domain knowledge should be used. But both require iterative measurement from individual ones.

Some error of Time series data may have some structure, which we can use it for making the estimation efficient by GLS method. We can find that structure of error by 2LSE, which we fit model on data and make model on residue again. Then the letter model can be used as weight like 2LSE, which we call it prewhitening.

Two least square estimation(2LSE) is a linear regression that use the residue for reducing bias. When we find some structure in residue, we can fit model for it and use the residue of it to adjust coefficients. Suppose we have a data matrix X and we want to reduce the dimension of X using weight matrix W and hidden node matrix H. And suppose we fit the model by $X = WH + E$

We know that OLS solutions for each W, H are $XH^T(HH^T)^{-1}, (W^TW)^{-1}W^TX$ and GLS solutions with weight matrix $V = EE^T, V' = E^TE$ are $XH^T\left(HV'^{-1}H^T\right)^{-1}, (W^TV^{-1}W)^{-1}W^TX$. These assume that errors are independent but there are some cases which that assumption is not satisfied. For examples, the individual data have some relationships with each others, which the group variable H can't explain. Then we can see E as individual effect which group H can' explain.

Then we can set model as $X = WH + R + E$ which R is pure individual effects. Then we can use 2LSE for estimating R if we make model residue E' which are derived from the model $X = WH$. Then we adjust W, H as $(XH^T - RH^T)(HV'^{-1}H^T)^{-1}, (W^TV^{-1}W)^{-1}(W^TX - W^TR)$ and use them for modifying R. We can iterate this algorithm if we need. Any method for reducing measurement error require proper assumption of the error. Or not, these make bias on coefficients. So we need to use domain knowledge for setting modelling.